



A sparse representation method for determining the optimal illumination directions in Photometric Stereo

Vasileios Argyriou^{b,*}, Stefanos Zafeiriou^a, Barbara Villarini^b, Maria Petrou^a

^a Communications and Signal Processing Research Group, Department of Electrical and Electronic Engineering, Imperial College, South Kensington Campus, London SW7 2AZ, UK

^b Faculty of Computing, Information Systems and Mathematics, Kingston University London, Penrhyn Road, Kingston-upon-Thames, Surrey KT1 2EE, UK

ARTICLE INFO

Article history:

Received 18 October 2012

Received in revised form

4 February 2013

Accepted 30 April 2013

Available online 9 May 2013

Keywords:

L1

Photometric stereo

3D imaging

Face reconstruction

ABSTRACT

The analysis of surface and texture details with the help of changes in illumination direction is a key task in 3D shape reconstruction either based on Photometric Stereo, Shape from Shading or Structured Light. This paper presents a novel approach for estimating the optimal illumination directions for the accurate calculation of the surface normals, while minimising the presence of shadows and the reconstructed albedo error. The method regards a sparse representation of the illumination arrangement and estimates the light directions using l_1 optimisation. The Lambertian model is considered and the theoretical development is demonstrated with experimental results.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The variation of the intensities observed in images depends on variation in both surface reflectance and surface relief. While the reflectance properties are intrinsic to a surface, the surface relief produces a pattern of shadings that depends strongly on the direction of illumination. The appearance of a 3D surface changes drastically with illumination. Different image details are enhanced for different illumination directions. In addition, depending on the imaging geometry, highlights may be created damaging the captured image. So, illumination and viewing directions play a crucial role in the quality of the produced image, and they should be carefully chosen for applications in which this is possible. This is more so for photometric stereo where estimates of local surface orientation and local surface albedo are obtained by using several images of the same surface taken from the same

viewpoint but illuminated from different directions. Sub-optimal geometric arrangements may crucially affect the reliability of the subsequently inferred information.

Woodham [40] was the first to introduce photometric stereo. He proposed a method which was simple and efficient, but only dealt with Lambertian surfaces and was sensitive to noise. In his method, the surface gradient can be recovered by using two photometric images, assuming that the surface albedo is already known for each point on the surface. Coleman and Jain [12] extended photometric stereo to four light sources, where specular reflections were discarded and estimation of surface shape could be performed by means of diffuse reflections and the use of the Lambertian model. Nayar et al. [28] developed a photometric approach which uses a linear combination of the Lambertian model and an impulse specular component to obtain the shape and reflectance information for a surface. Barsky and Petrou [3–5] presented an algorithm for estimating the local surface gradient and real albedo by using four source colour photometric stereo in the presence of highlights and shadows. It is also worth mentioning the related work presented in Solomon and Ikeuchi [33],

* Corresponding author.

E-mail address: Vasileios.Argyriou@kingston.ac.uk (V. Argyriou).

Chandraker et al. [10], Levine and Bhattacharyya [25], Finlayson et al. [15], Smith and Hancock [32], Sun et al. [36], Alldrin et al. [1], Hertzmann and Seitz [20], Georgiades [17], Tagare and deFigueiredo [37], Ragheb and Hancock [29], Argyriou et al. [38], Argyriou and Petrou [39], Zhang et al. [44], Fyffe et al. [16], Sakaue and Sato [31], and Miyazaki and Ikeuchi [26].

In this paper, we analyse the problem of estimating the optimal imaging configuration using a sparse representation method. In Section 2 we briefly present some previous work on the subject. Standard photometric stereo is reviewed in Section 3. In Section 4 our methodology on optimal illumination directions estimation is presented. Experiments are presented in Section 5 and conclusions in Section 6.

2. Previous work

The problem of estimating the proper lighting arrangement in 3D surface reconstruction methodologies based on photometric stereo has been considered in the past [41,24,34,11]. Both the number of required light sources and the optimal illumination configurations in terms of azimuth φ_l and zenith θ_l angles of each light source l have been evaluated and studied in order to provide more accurate normals and 3D reconstructions. The use of dense iso-intensity contours was recommended by Woodham [41] to obtain maximum accuracy, since in this case a small change in the surface gradient components p_n and q_n results in a large intensity change. In order to achieve dense iso-intensity contours the zenith angle θ_l is increased, but this increases the number of shadows, which limits the effectiveness of the algorithm. Regarding the azimuth angles of the light sources, Woodham pointed out that the illumination vectors must not be co-planar, otherwise the illumination matrix \mathbf{L} is not invertible. The illumination matrix is made up from the directions of the different illuminants written as rows. If $\mathbf{n}(x,y)$ is the normal vector for surface patch (x,y) , $\rho_S(x,y)$ is the albedo of the same patch and $\mathbf{A}_g(x,y)$ is the vector representing the grey values in the corresponding images for this patch, we have

$$\mathbf{A}_g(x,y) = \rho_S(x,y)\mathbf{L}\mathbf{n}(x,y) \quad (1)$$

In the case of a two image photometric stereo, Lee and Kuo [24] deduced that it is desirable to incorporate reflectance maps that compensate each other's weaknesses, in order to determine the optimal illumination configuration. Observing that the azimuth angle φ_l of the illumination vector determines the orientation of the reflectance map about the origin of the axes, whereas the zenith angle θ_l determines the distance between the origin of the axes and the point of maximum illumination, i.e. the point that has normal vector parallel to the vector towards the light source, determined by coordinates (p_l, q_l) in the gradient space, the angular difference between two reflectance maps would be given by $|\varphi_{l_1} - \varphi_{l_2}|$.

Gullon [18] confirmed that the two image photometric stereo is more sensitive to the azimuth rather than the zenith angle difference and that the optimal value for the difference in the azimuth angles of the two lighting

directions is 90° . Furthermore, Gullon suggested that distributing the illumination azimuth angles uniformly through 360° is optimal when three-light photometric stereo is considered. A theoretical analysis of Gullon's arrangement was presented by Spence and Chantler [34,35] based on the sensitivity analysis of photometric stereo deriving expressions of each surface normal vector with respect to image intensities.

Due to the dependance on the surface shape and its statistics, it was found in Spence and Chantler [34] that the optimal azimuth and zenith angles cannot be specified and that the configuration that results in the minimum noise is not unique. In the case of the common zenith angle being constrained, the optimal values for azimuth angles were estimated and it was suggested to use 120° angle difference in a three-image Lambertian photometric stereo configuration. This result is in agreement with the work of Gullon [18], where a uniform distribution of illumination directions was recommended.

The optimal zenith angle in case of uniformly distributed light sources, according to the azimuth angle, has been found to be around 55° , but if shadows are present the angle should be reduced [34,18]. On the contrary, if the surface is smooth and shadows are not an issue, the zenith angle can be increased. Furthermore, Drbohlav and Chantler [14] extended the above for n light sources and deduced the same optimal zenith angle when the sources were equally spaced in azimuth angles of $360/n$ degrees.

Regarding the number of light sources required, Coleman and Jain [12], Solomon and Ikeuchi [33], Barsky and Petrou [4,5] and Chandraker et al. [10] proposed methodologies requiring four light sources. Rushmeier et al. [30] proposed a five light source photometric stereo system, while a six light source photometric stereo technique was suggested by Sun et al. [36] employing a slightly more sophisticated decision criterion so as to discard pixels with doubtful values.

In this paper the optimal illumination configurations in terms of azimuth φ_l and zenith θ_l angles have been worked out using l_1 optimisation of a criterion function defined in Section 4. The proposed methodology is evaluated using standard four light photometric stereo but it may be generalised and combined with any other photometric stereo algorithm.

3. Photometric stereo for Lambertian surfaces

For a Lambertian object illuminated by a light source of parallel rays, the observed image intensity \mathbf{a} at each pixel is given by the product of the albedo ρ and the cosine of the incidence angle θ_i (the angle between the direction of the incident light and the surface normal) [21]. The above incidence angle can be expressed as the dot product of two unit vectors, the light direction \mathbf{l} and the surface normal \mathbf{n} , $\mathbf{a} = \rho \cos(\theta_i) = \rho(\mathbf{l} \cdot \mathbf{n})$.

Let us now consider a Lambertian surface patch with albedo ρ and normal \mathbf{n} , illuminated in turn by several fixed and known illumination sources with directions $\mathbf{l}^1, \mathbf{l}^2, \dots, \mathbf{l}^Q$, where Q is the total number of light sources. In this case we can express the intensities of the obtained

(grey scale) pixels as

$$\mathbf{a}_k = \rho(\mathbf{l}_k \cdot \mathbf{n}), \quad \text{where } k = 1, 2, \dots, Q. \quad (2)$$

We stack the pixel intensities to obtain the pixel intensity vector $\mathbf{A}_a = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_Q)^T$. Also the illumination vectors are stacked row-wise to form the illumination matrix $\mathbf{L} = (\mathbf{l}^1, \mathbf{l}^2, \dots, \mathbf{l}^Q)^T$. Eq. (2) could then be rewritten in matrix form

$$\mathbf{A}_a = \rho \mathbf{L} \mathbf{n} \quad (3)$$

If there are at least three illumination vectors which are not coplanar, we can calculate ρ and \mathbf{n} using the Least Squares Error technique, which consists of using the transpose of \mathbf{L} , given that \mathbf{L} is not a square matrix:

$$\mathbf{L}^T \mathbf{A}_a = \rho \mathbf{L}^T \mathbf{L} \mathbf{n} \Rightarrow (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{A}_a = \rho \mathbf{n} \quad (4)$$

Since \mathbf{n} has unit length, we can estimate both the surface normal (as the direction of the obtained vector) and the albedo (as its length). Extra images allow one to recover the surface parameters more robustly.

4. Finding the optimal light positions using

l_1 optimisation

In this section we describe a method for finding the optimal light position for a given type of 3D surface, by trying to eliminate the shadows while recovering the albedo of each facet of the surface. From these lights we shall select the Q lights with which we need to perform Photometric Stereo (PS). Usually $Q < 8$, which is also adequate for simulation application using OpenGL or other rendering tools increasing simultaneously the overall performance.

In other words, in the optimisation stage, we try to find the smallest number m of required lights for both the exclusion of all shadows and for albedo reconstruction. In case $m > 8$, we select the best 8 of them and consider all their Q -light combinations, with the restriction of not being collinear. Then the best combination is selected. The term “best” in this context will be defined later, after Eq. (10). In case $Q < m < 8$, again all combinations of m by Q are tested and the best one is selected. In case $m < Q$, extra lights are added, selected as the “best” from among those that were not picked by the optimisation algorithm, because they did not fulfill some of the criteria used by that algorithm. This is so that we always have a predefined fixed number of lights Q . In our scenarios $Q=4$. It should be mentioned at this stage that not only the performance in terms of accuracy increases with the number of lights Q but also the cost and the computational complexity. The actual improvement in performance depends on the particular shape of the reconstructed scene and regarding the proposed approach it will also improve the reconstruction independent of the characteristics of the observed shapes.

The proposed method does not make any assumptions about the underlying illumination model. However, we assume that we know the generic shape of the surfaces that are to be reconstructed by photometric stereo. So, the idea is to work out the optimal illumination directions for the particular type of surface and then use them for all subsequent surfaces. For example, if we want to

reconstruct faces, we use a generic model face to work out the optimal light arrangement, which may subsequently be used for all real faces we wish to image. Regarding the number of samples it depends on the particular object class since they are required to generate an average or generic surface. Another approach would be to utilise a prototype to obtain the generic surface and in that case only one actual object is required. About the complexity of the optimisation stage it is not an issue and does not affect the performance of the system since it is performed only once per object class and not during the actual reconstruction. Also in the case a generic shape is not possible to be extracted, a recursive approach could be introduced using an initial setup (e.g. [14]) and then based on the obtained initial reconstruction the optimal illumination directions could be estimated using the same approach.

In order to develop the proposed method we make use of recent advances in the sparse representation theory. According to the sparse representation theory, sparse signals can be exactly reconstructed from a small number of measurements [13,9,8,23]. These principles have been used for face recognition, image super-resolution and face hallucination [42]. Wright et al. [42] motivated by the principles of compressed sensing, tried to represent an object using a sparse linear combination of an overcomplete dictionary. In particular, a facial image was represented as a sparse linear combination of the training facial images. It was shown that, when a sufficient number of training samples were available from each facial class, it was possible to represent the test samples as linear combinations of just the training samples from the same facial class. The resulting optimisation problem penalised the l_1 -norm of the coefficients in the linear combination. The authors showed that the representation was indeed sparse, involving only a small fraction of the overall training database. They also argued that the calculation of the sparsest representation is a way of performing discriminant analysis between the facial classes. This intuitively means that the test image is most likely to belong to the facial class with the most nonzero coefficients.

Let us consider all possible positions of lights uniformly distributed on the surface of a hemisphere, with the inspected object located at the center of the hemisphere and the camera at the zenith of the object. Moreover, let us assume that a light exists in each position, and let us use these lights to acquire images from a 3D object, that is a generic representative of the type of surface we wish to inspect. For example, if we wish to inspect human faces, this object could be a properly painted mask, or a typical face. For every possible light corresponds an image and a shadow map (i.e. a map marking the pixels that are turned away from the illuminating source). An example of the above is shown in Fig. 1. We assume that in order to reconstruct the original albedo of the surface and to eliminate the shadows (i.e. obtain a shadowless surface) we only need a limited number of lights which should be specific for the 3D object in hand. So, we have to identify from among all possible positions of the lights on the hemisphere of Fig. 1, the subset that is adequate for our

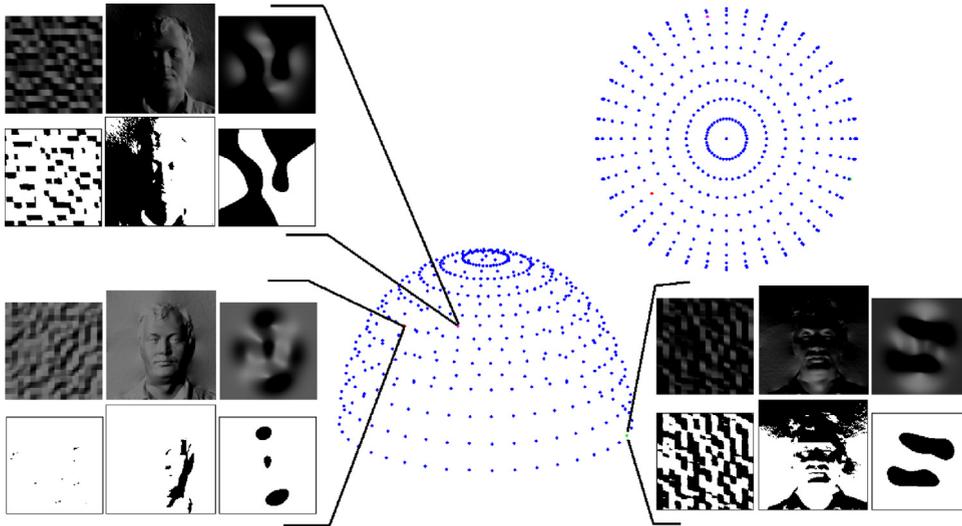


Fig. 1. The hemisphere structure and the sampled lights. For each light and each 3D surface correspond a grey scale image and a shadow map, respectively.

problem. In other words, we formulate the problem of finding the optimal directions of lights for a given object as the search of the sparsest set of lights in order to reconstruct the albedo and a ‘shadow map’ without any shadowed parts (Fig. 2).

Let \mathbf{a}_i and \mathbf{s}_i be the grey scale image and the shadow map that are derived from the i -th light, respectively. Let also \mathbf{a}_o be the albedo of the reference surface used for training and \mathbf{s}_o the shadow map with the minimum number of shadows.

We now create the dictionary $\mathbf{A}_a = [\mathbf{a}_1 | \dots | \mathbf{a}_N]$, which is a matrix with columns of the images captured under all N -sample lights and the dictionary $\mathbf{A}_s = [\mathbf{s}_1 | \dots | \mathbf{s}_N]$, which contains the shadow maps for all the sample lights. We shall first consider the two components, i.e. the images and the shadow maps, separately and then we shall propose a fusing scheme.

In the case of the images, we seek to find the sparsest vector \mathbf{w}_a such that the albedo \mathbf{a} can be written as a linear combination of the columns of dictionary \mathbf{A}_a . This optimisation problem then is as follows:

$$\tilde{\mathbf{w}}_a^0 = \arg \min \|\mathbf{w}_a\|_0 \quad \text{subject to } \mathbf{A}_a \mathbf{w}_a = \mathbf{a} \quad (5)$$

where $\|\cdot\|_0$ denotes the l_0 -norm, which counts the number of nonzero entries in a vector. Unfortunately, the problem of finding the sparsest solution $\tilde{\mathbf{w}}_a^0$ is NP-hard, and difficult to solve even approximately. Recent developments in the emerging theory of sparse representations and compressed sensing [13,8] reveal that if the sought solution \mathbf{w}_a^0 is sparse enough, the solution of the l_0 -minimisation problem (5) is equal to the solution of the following l_1 -minimisation problem:

$$\tilde{\mathbf{w}}_a^1 = \arg \min \|\mathbf{w}_a\|_1 \quad \text{subject to } \mathbf{A}_a \mathbf{w}_a = \mathbf{a}. \quad (6)$$

Alternatively, if we allow an error in the reconstruction, the problem becomes

$$\tilde{\mathbf{w}}_a^1 = \arg \min \|\mathbf{w}_a\|_1 \quad \text{subject to } \|\mathbf{A}_a \mathbf{w}_a - \mathbf{a}\|_2 < \epsilon_a. \quad (7)$$

Usually $\epsilon_a \sim 10^{-3}$. This problem can be solved in polynomial time using the algorithm presented in Candes and Romberg [7].

In a similar fashion, we can use the shadow maps \mathbf{s}_i for all the lights $i=1, \dots, N$ in order to build the dictionary and try to find a sparse vector \mathbf{w}_s so that we have a shadow map \mathbf{s} with minimum number of shadows:

$$\tilde{\mathbf{w}}_s^1 = \arg \min \|\mathbf{w}_s\|_1 \quad \text{subject to } \|\mathbf{A}_s \mathbf{w}_s - \mathbf{s}\|_2 < \epsilon_s. \quad (8)$$

Now, let us try to find the sparse vector of lights \mathbf{w} using both albedo and shadow map information, that is, for the same vector \mathbf{w} having both $\mathbf{A}_s \mathbf{w}_s = \mathbf{s}$ and $\mathbf{A}_a \mathbf{w}_a = \mathbf{a}$. The optimisation problem can be formally written as

$$\begin{aligned} \tilde{\mathbf{w}}^1 = \arg \min \|\mathbf{w}\|_1 \\ \text{subject to } \|\mathbf{A}_s \mathbf{w} - \mathbf{s}\|_2 < \epsilon_a \\ \text{and } \|\mathbf{A}_a \mathbf{w} - \mathbf{a}\|_2 < \epsilon_s. \end{aligned} \quad (9)$$

We shall try to solve the above optimisation problem using a different formulation. Let us create the composite concatenated vector \mathbf{f}_c containing both the image and the shadow map of all sampled lights $\mathbf{f}_c = \begin{bmatrix} \mathbf{a} \\ \mathbf{s} \end{bmatrix}$ and the concatenated dictionary \mathbf{A}_c containing the albedos and the shadow maps of all sampled lights $\mathbf{A}_c = \begin{bmatrix} \mathbf{A}_a \\ \mathbf{A}_s \end{bmatrix}$. Then let us try to identify vector \mathbf{w}_c , such that

$$\tilde{\mathbf{w}}_c^1 = \arg \min \|\mathbf{w}_c\|_1 \quad \text{subject to } \|\mathbf{A}_c \mathbf{w}_c - \mathbf{f}_c\|_2 < \epsilon_c. \quad (10)$$

After the calculation of a sparse vector \mathbf{w}^1 from one of the optimisation problems (7), (8) or (10), we choose the best subset \mathcal{I} of lights (usually between 4 and 8), i.e. the subset of m lights with the largest coefficients w_i . So, if we want to select Q -lights we take all possible Q -light combinations $\binom{m}{Q}$ if there are m nonzero entries in \mathbf{w}^1 . Let us now assume that we have in total C_n sets \mathcal{I}_i with each of them containing Q -lights. Let $\mathbf{w}^1(\mathcal{I}_i)$ denote vector:

$$[\mathbf{w}^1(\mathcal{I}_i)]_j \triangleq \begin{cases} [\mathbf{w}^1]_j & \text{if } j \in \mathcal{I}_i \\ 0 & \text{if } j \notin \mathcal{I}_i. \end{cases} \quad (11)$$

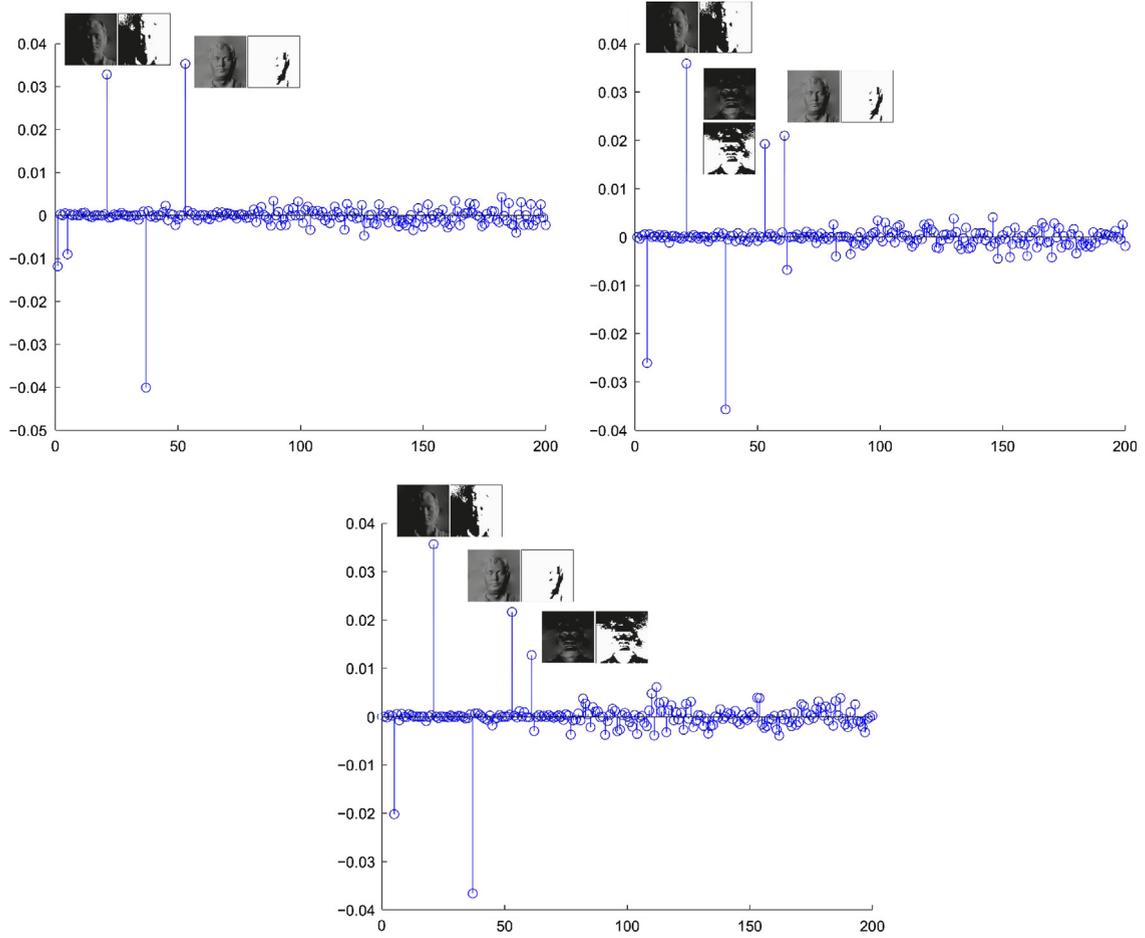


Fig. 2. An example of sparse decomposition of the shadowless image and albedo. Along the vertical axis we measure the weight by which the corresponding image is multiplied and combined with the others to yield the albedo of the surface, and along the horizontal axis we list all the illumination directions.

For all sets \mathcal{I}_i , calculate the corresponding residual. For example, for the set coming from the solution of optimisation problem (7), we compute

$$r(\mathcal{I}_i) \triangleq \|\mathbf{a} - \mathbf{A}_a \mathbf{w}^1(\mathcal{I}_i)\|_2 \quad (12)$$

Then, among all the different sets \mathcal{I}_i , we choose the one with minimum residual:

$$\mathcal{I}_{opt} = \arg \min_{\mathcal{I}_i} r(\mathcal{I}_i) \quad (13)$$

The above algorithm can be summarised as follows.

Algorithm 1.

- Data** Arrange N illuminants to lie on the surface of a hemisphere.
- initialisation**
- foreach** *light* **do**
 - calculate a shadow map $\mathbf{s}_i \in [0, 1]^{M_1 \times M_2}$;
 - calculate the intensity image $\mathbf{a}_i \in [0, 255]^{M_1 \times M_2}$ of the 3D reference object used for training;
- end**
- Result:** Downsample, or project using orthogonal random bases, the shadow map to $\mathbf{s}_i \in \mathbb{R}^p$ and the image to $\mathbf{a}_i \in \mathbb{R}^p$ and create the dictionaries \mathbf{A}_s and \mathbf{A}_a .
- Result:** Solve one of the optimisation problems (7), (8) or (10) in order to derive a sparse set of weights \mathbf{w}^1 for the illuminants.

Result: From the set \mathbf{w}^1 choose a set of $m \ll N$ lights (in our case $m=8$) with the largest coefficients such that every four of them are linearly independent.

forall n -combinations $\mathcal{I}_1, \dots, \mathcal{I}_n$ (in our case $n = \binom{m}{4}$) **do**
 Calculate the residual using (12) or the corresponding equation, according to which optimisation problem you have solved, and choose the one with the minimum \mathcal{I}_{opt} (Eq. (13)).

end

5. Experiments and results

In our experiments three sets of simulated surfaces were used for evaluation, namely isotropic surfaces, anisotropic surfaces and simulated human faces. Additionally, for each surface, three different textures were used to simulate albedo. In the case of human faces a uniform grey colour was used, as well as real skin albedo. For surface reconstruction, the standard four light photometric stereo was used, without applying any sophisticated algorithms to discard shadows or highlights. Also it should be mentioned that other methods could be used for the evaluation, such as the work of Wu et al. [43], Hernandez et al. [19], and Hyeongwoo et al. [22]. For each set of surfaces

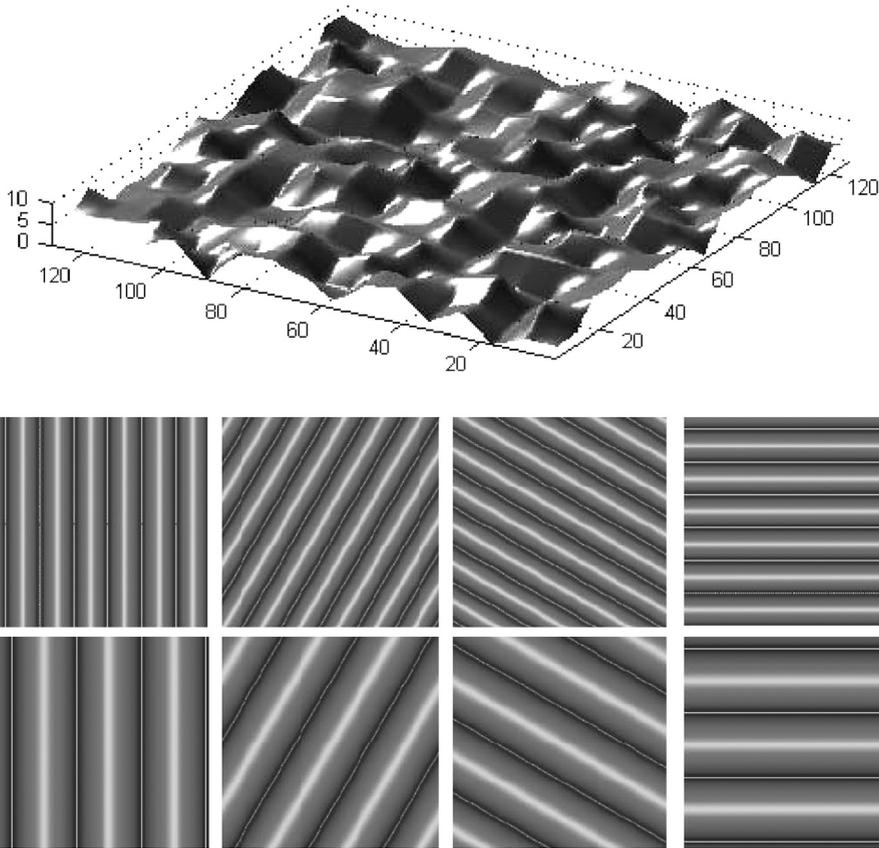


Fig. 3. An example of a simulated isotropic surface of size 128×128 pixels; and two example surfaces from each of the four sets of anisotropic surfaces used in our experiments.

four light sources were obtained using the proposed methodology. A subset of each surface class was used to obtain the optimal illumination directions for the corresponding set. From each surface in the training subset the optimal illumination directions were identified and the corresponding vectors were averaged to yield a single set of optimal illumination directions for the particular class of surface. Finally, we performed experiments with real objects, where the ground truth was obtained using the 3dMD 3D surface imaging system [27].

The three different proposed methodologies to obtain the optimal illumination configuration, i.e. considering only the shadow maps, only the albedo, or their combination, were compared with the configuration proposed by Drbohlav and Chantler [14]. According to Drbohlav and Chantler [14] the zenith angle should be the same for all light sources and equal to 55° , while the azimuth directions should be equally spaced pointing at the four corners of a square.

Since simulated data were used and the real surface normals were available, the Lambertian model was applied to generate four images of each surface, corresponding to the estimated optimal light sources. Standard photometric stereo was then applied on these images to obtain the surface normals. In addition, experiments were performed with uniform and non-uniform albedos.

In order to compare the performance of the proposed approaches, the angular error (AE) measure suggested by

Barron et al. [2] was used:

$$\psi_{AE} = \cos^{-1} \left[\frac{x_e x_c + y_e y_c + z_e z_c + 1}{\sqrt{1 + x_e^2 + y_e^2 + z_e^2} \sqrt{1 + x_c^2 + y_c^2 + z_c^2}} \right] \quad (14)$$

where $(x_e, y_e, z_e)^T$ and $(x_r, y_r, z_r)^T$ are the estimated and the real surface normals, respectively.

Furthermore, since the real albedo was available, the mean absolute difference was used to compare the performance of the proposed methodologies

$$\epsilon_{AD} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |A_e(i, j) - A_r(i, j)| \quad (15)$$

where A_e and A_r are the estimated and the real albedos, respectively.

Finally, at the last part of the evaluation, where experiments were performed under real environmental conditions, the sum of the absolute height map difference, between the estimated reconstructed surface, using Photometric Stereo, and the ground truth, obtained using the 3dMD imaging system, was used to compare the accuracy of the proposed methodologies:

$$h_{EG} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |H_e(i, j) - H_r(i, j)| \quad (16)$$

Here H_e and H_r are the estimated and the real height maps, respectively.

5.1. Experiments using isotropic surfaces

Experiments were performed with the first set of simulated data, where fifteen isotropic surfaces were selected for training and fifteen for evaluation (see Fig. 3). The surfaces were containing random peaks with low altitude. In Fig. 4 the performances of the three proposed methodologies using shadow maps, the albedo and their combination are evaluated, against the ‘default’ illumination configuration [14], in terms of angular and mean absolute difference. The performances of the proposed algorithms are identical in that scenario and observing the results we may see that the proposed methods

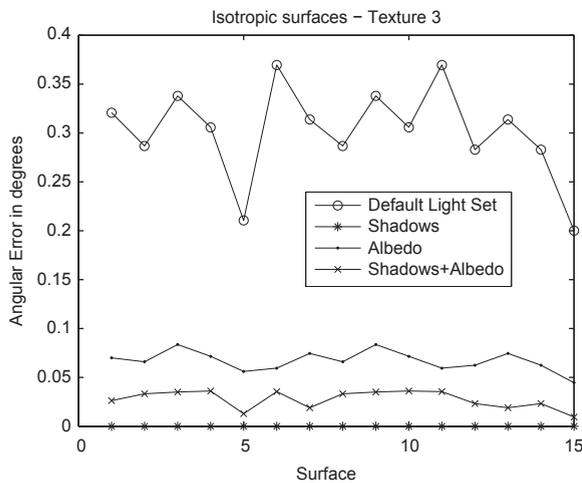


Fig. 4. Performance comparison of the proposed methodologies using shadow maps, albedo and their combination, evaluated in terms of mean angular error for the recovered orientation, for all isotropic surfaces.

Table 1

The mean angular error (MAE) computed over all simulated surfaces and for all methodologies (i.e. S for shadow maps, A for albedo and S+A for both of them). The best result for each case is in bold.

Surfaces	Default	S	A	S+A
Isotropic	0.3012	0.0000	0.1724	0.0000
Anisotropic 0°	2.5334	0.1842	0.6300	1.1778
Anisotropic 45°	3.0619	0.5611	0.9842	2.2823
Anisotropic 135°	3.0612	0.2546	1.3338	2.6132
Anisotropic 90°	2.5692	0.2133	0.2299	0.0135
Faces	3.4566	0.1949	3.1321	2.8468

Table 2

The mean absolute difference (MAD) computed over all 15 isotropic simulated surfaces for different textures and for all methodologies (i.e. S for shadow maps, A for albedo and S+A for both of them). The best result for each case is in bold.

Method	Texture 1	Texture 2	Texture 3
Default	0.00023	0.00026	0.00017
S	0.00000	0.00000	0.00000
A	0.00009	0.00023	0.00022
S+A	0.00000	0.00000	0.00010

improve significantly the accuracy of the estimated surface normals and the albedo obtained by using the default illumination arrangement (i.e. the one that is optimal for any type of surface when we have not any prior knowledge about it). In Tables 1 and 2 the mean angular error (MAE) and the mean albedo errors (MAD) are shown for all textures, respectively.

5.2. Experiments using anisotropic surfaces

Next, four sets of seven anisotropic surfaces (Fig. 3) were used. For each set, a subset of three surfaces was selected for training and the remaining four for evaluation. The same experiments were performed and the results for the three proposed methodologies are shown in Fig. 5. In terms of the angular error, the performance of the proposed methods based on shadows only and on shadows plus the albedo is

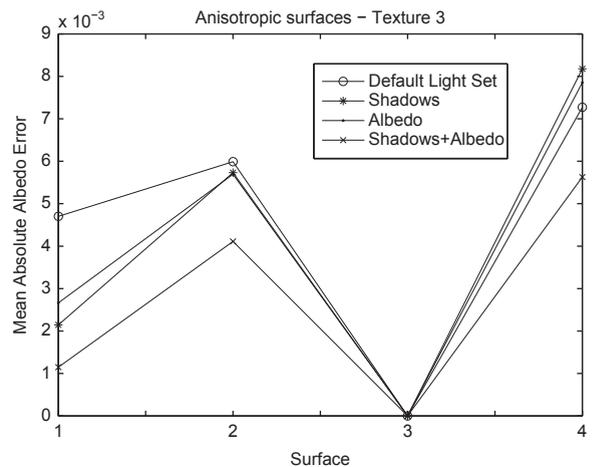


Fig. 5. Performance comparison of the three proposed methodologies, using shadow maps, the albedo and their combination, evaluated in terms of mean absolute error of the albedo of the recovered surface, for anisotropic surfaces.

Table 3

The mean absolute difference (MAD), between the true and the recovered normal field, computed over all anisotropic simulated surfaces for different textures, for all methodologies (i.e. S for shadow maps, A for albedo and S+A for both of them) and for all illumination configurations. The best result for each case is in bold.

MAD	0°			45°		
	T 1	T 2	T 3	T 1	T 2	T 3
Default	0.0047	0.0052	0.0034	0.0096	0.0107	0.0070
S	0.0019	0.0034	0.0027	0.0071	0.0031	0.0037
A	0.0045	0.0013	0.0019	0.0090	0.0105	0.0058
S+A	0.0066	0.0079	0.0017	0.0119	0.0148	0.0080
	135°			90°		
	T 1	T 2	T 3	T 1	T 2	T 3
Default	0.0096	0.0107	0.0070	0.0048	0.0053	0.0035
S	0.0039	0.0074	0.0051	0.0025	0.0021	0.0010
A	0.0092	0.0105	0.0073	0.0024	0.0035	0.0028
S+A	0.0174	0.0189	0.0127	0.0005	0.0076	0.0019

identical. The average estimates over all anisotropic surfaces are displayed in Tables 1 and 3, indicating further that the proposed methodology provides better results due to the more accurate illumination configuration, designed taking into consideration the prior knowledge concerning the type of surface that is to be reconstructed.

5.3. Experiments using faces

Further experiments were performed using faces (Fig. 6). The illumination configuration using the default setup is compared with the configurations obtained by using the training set of faces. Fig. 7 shows the performance of the proposed methodologies both using the angular difference, to assess the accuracy of the recovered normal field and albedo field, for the test faces. All the mean errors are shown in Tables 1 and 4.

5.4. Experiments under real conditions

Having learnt the optimal illumination directions from the simulated faces and anisotropic surfaces, now we are going to test these arrangements with real surfaces of similar type, for which the ground truth is available with the help of the 3dMD imaging system. In particular, two mannequin faces will be reconstructed using the optimal illumination directions reported in Tables 5, 6 and 7 and the reconstruction will be compared with the real height maps worked out by the 3dMD scanner. In addition, a bottle, which may be thought of as an anisotropic surface with vertical ribs, will also be tested using the optimal illumination directions for such surfaces also reported in Tables 5, 6 and 7. Some of them are demonstrated in Fig. 8.

Each object was illuminated both from the default and the estimated optimal directions. Using the images captured with the default and the estimated optimal light sources, Photometric Stereo and integration were applied in succession, in order to obtain the 3D surfaces for both illumination configurations. Using the Iterative Closest Point algorithm [6] the obtained surfaces are aligned with the corresponding surfaces obtained from the 3dMD imaging system. In order to evaluate the accuracy of the

obtained surfaces, the absolute height map difference defined by Eq. (16) was used. The reconstructed 3D surfaces for both configurations are shown in Fig. 9. The average absolute height map difference is reported in Table 8 with the proposed illumination configuration resulting in the least error.

The proposed algorithm was further applied to the reconstruction of three real human faces (see Fig. 10) using photometric data captured both with the default and the

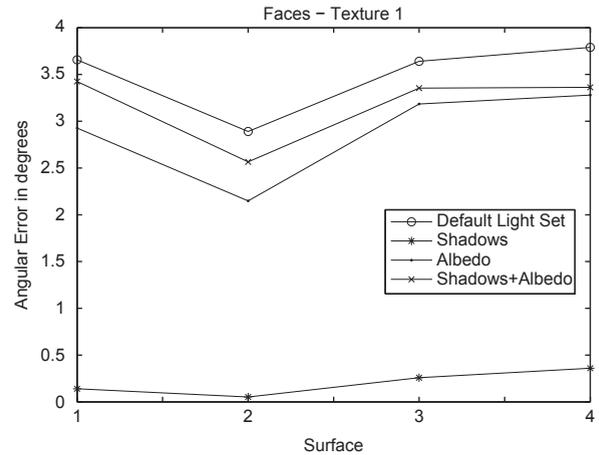


Fig. 7. Performance comparison of the three proposed methodologies, using shadow maps, albedo and their combination, evaluated in terms of the mean angular error of the recovered field of normals for the simulated face images.

Table 4

The mean absolute difference (MAD), between the true and the normal field of the recovered surface, computed over all 4 simulated faces for a uniform grey colour and for the mean albedo, and for all methodologies (i.e. S for shadow maps, A for albedo and S+A for both of them). The best result for each case is in bold.

Method	Grey	Skin
Default	0.0092	0.0081
S	0.0025	0.0001
A	0.0075	0.0012
S+A	0.0102	0.0116



Fig. 6. Examples of simulated faces used in our experiments. The faces at the top row were used for training and the faces at the bottom for testing.

Table 5The default and the proposed illumination directions using the optimisation step only *shadow maps*, for faces, isotropic and anisotropic surfaces.

Light	$\varphi_{l_1}, \theta_{l_1}$	$\varphi_{l_2}, \theta_{l_2}$	$\varphi_{l_3}, \theta_{l_3}$	$\varphi_{l_4}, \theta_{l_4}$
Default	(45°, 55°)	(135°, 55°)	(315°, 55°)	(225°, 55°)
Isot	(43°, 37°)	(136°, 36°)	(304°, 32°)	(225°, 33°)
Anisot 0°	(50°, 38°)	(129°, 37°)	(316°, 42°)	(228°, 34°)
Anisot 45°	(17°, 33°)	(119°, 42°)	(316°, 26°)	(215°, 46°)
Anisot 135°	(42°, 24°)	(153°, 28°)	(334°, 33°)	(215°, 29°)
Anisot 90°	(63°, 37°)	(147°, 28°)	(334°, 33°)	(226°, 41°)
Faces	(48°, 23°)	(122°, 29°)	(309°, 32°)	(211°, 29°)

Table 6The default and the proposed illumination directions using the optimisation step only the *albedo*, for faces, isotropic and anisotropic surfaces.

Light	$\varphi_{l_1}, \theta_{l_1}$	$\varphi_{l_2}, \theta_{l_2}$	$\varphi_{l_3}, \theta_{l_3}$	$\varphi_{l_4}, \theta_{l_4}$
Default	(45°, 55°)	(135°, 55°)	(315°, 55°)	(225°, 55°)
Isot	(42°, 36°)	(145°, 47°)	(318°, 51°)	(222°, 38°)
Anisot 0°	(19°, 49°)	(117°, 52°)	(296°, 43°)	(194°, 59°)
Anisot 45°	(43°, 33°)	(140°, 26°)	(318°, 24°)	(212°, 48°)
Anisot 135°	(73°, 45°)	(151°, 45°)	(319°, 42°)	(233°, 21°)
Anisot 90°	(47°, 32°)	(144°, 42°)	(335°, 35°)	(230°, 33°)
Faces	(65°, 35°)	(151°, 59°)	(306°, 17°)	(202°, 63°)

Table 7The default and the proposed illumination directions using the optimisation step both *shadow maps* and *albedo*, for faces, isotropic and anisotropic surfaces.

Light	$\varphi_{l_1}, \theta_{l_1}$	$\varphi_{l_2}, \theta_{l_2}$	$\varphi_{l_3}, \theta_{l_3}$	$\varphi_{l_4}, \theta_{l_4}$
Default	(45°, 55°)	(135°, 55°)	(315°, 55°)	(225°, 55°)
Isot	(36°, 28°)	(151°, 49°)	(320°, 47°)	(209°, 29°)
Anisot 0°	(27°, 22°)	(132°, 39°)	(295°, 32°)	(186°, 35°)
Anisot 45°	(75°, 32°)	(135°, 35°)	(306°, 29°)	(222°, 46°)
Anisot 135°	(64°, 42°)	(145°, 54°)	(324°, 56°)	(200°, 31°)
Anisot 90°	(48°, 33°)	(153°, 40°)	(317°, 31°)	(239°, 37°)
Faces	(55°, 47°)	(157°, 50°)	(308°, 51°)	(211°, 61°)

proposed illumination directions. The person is assumed to be still during the acquisition stage, since a high speed camera was used for the acquisition (i.e. 200 frames per second), eliminating the registration problem.

In Fig. 11 results of the reconstructed faces obtained from the two compared illumination setups are shown. Observing the results it can be inferred that the proposed illumination directions result in more accurate estimates especially at the regions where moles or hair are present indicating that the proposed illumination directions provide more accurate and detailed reconstructions for faces. Furthermore, the side view was used to evaluate the reconstructed faces. The background was extracted manually and the Hausdorff distance was used to compare the reconstructions with the original profiles. Table 9 shows the results for all the faces.

6. Discussion and conclusions

In this paper, a method for estimating the optimal illumination configuration based on L_1 optimisation was

presented, for use with photometric stereo. Observing the results we may say that the proposed methodologies provide better illumination configurations that adapt to the class shape characteristics. Furthermore, depending on the characteristics of the surface and the albedo, it may be the solution of the optimisation problem using either the shadow maps or the albedo that produce the best performance. A training set of surfaces from each class is required for all methodologies. This may be obtained by either using laser scanners or any other 3D reconstruction technique. Regarding the case of the albedo approach, it may operate without requiring a training set, if we assume without significant error that the albedo corresponds to the brightness of the average image. Prior knowledge concerning the statistical distribution of the facets of the surface to be reconstructed was utilised. Such information may be available if we know the class of objects that is to be inspected. Experiments with simulated and real surfaces were performed in order to evaluate the performance of the proposed scheme. The mean angular error

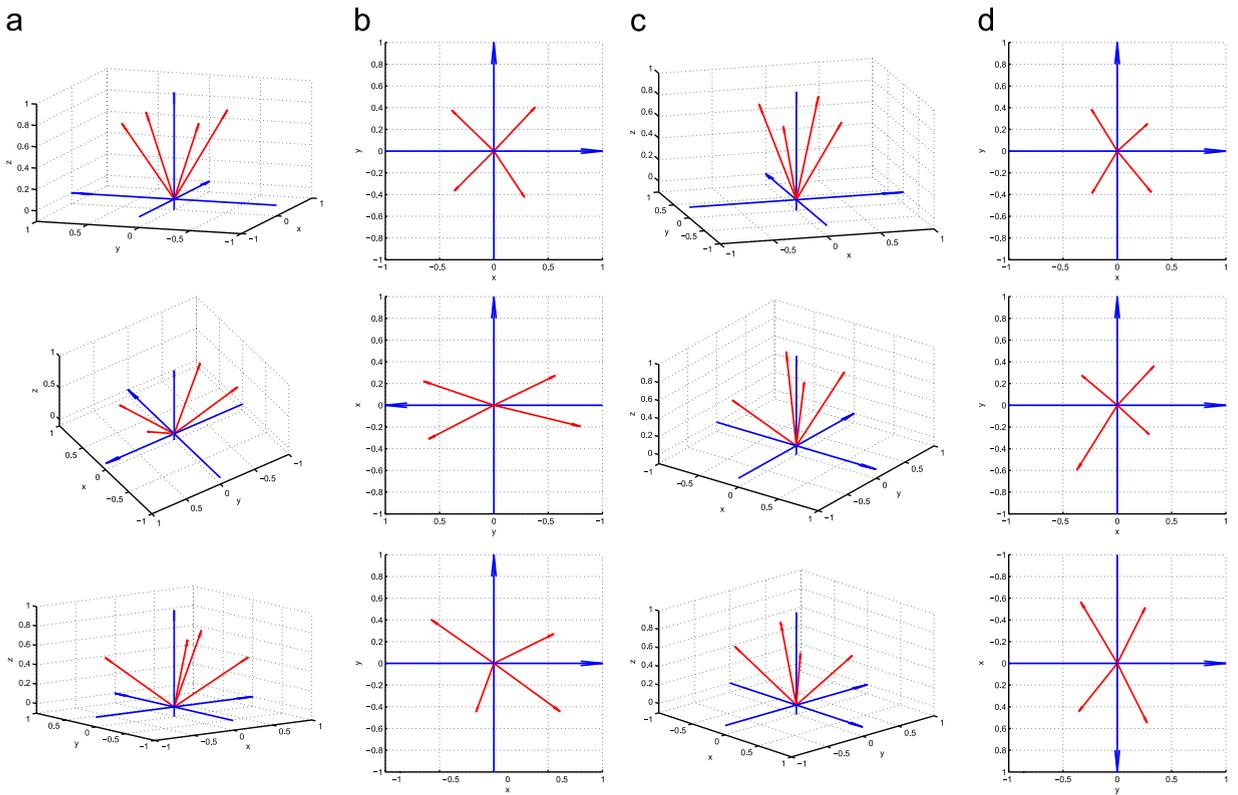


Fig. 8. The illumination directions in 3D representation (columns (a) and (c)) and in 2D top view (columns (b) and (d)) for isotropic surfaces (top left), faces (top right), an anisotropic surface 0° (mid left), an anisotropic surface 135° (mid right), an anisotropic 45° (bottom left) and an anisotropic 90° (bottom right).

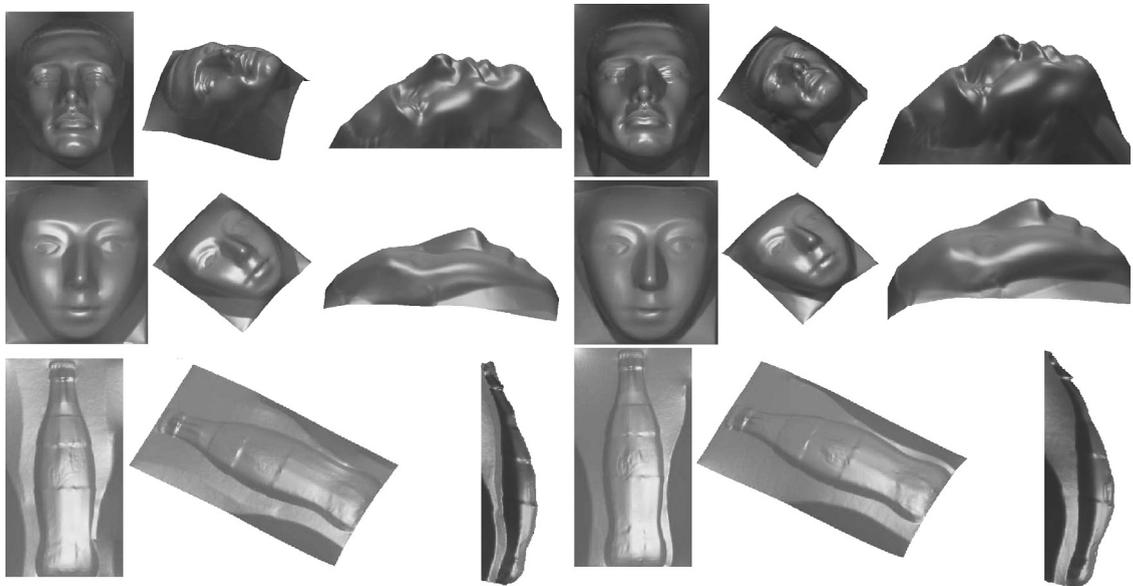


Fig. 9. 3D surfaces obtained using the default (left) and the proposed (right) illumination configuration for ‘Adam’, ‘Eve’ and ‘bottle’.

of the surface normals and the mean albedo difference were used to evaluate the performance of the proposed methodology in comparison with the illumination configuration proposed by Drbohlav and Chantler [14]. A

standard 4-lights photometric stereo, without applying any steps to eliminate errors due to shadows, was used to obtain the surface normals and the albedo. From the results it could be inferred that the proposed approach

Table 8

The mean absolute height map difference (MAD) for the recovered normal field, computed over all three faces captured under real environmental conditions using the default and the proposed optimal illumination configuration. The best result for each case is in bold.

Method	Adam	Eve	bottle
Default	69.4904	26.7044	17.0176
Proposed	59.3856	23.2544	13.7375

Table 9

The Hausdorff distance of the side views of the faces from the reconstructed side views for the tested illumination directions. The best result for each face is in bold.

Method	Face A	Face B	Face C
Default	54.3386	54.7264	48.4010
Proposed	45.8947	40.6348	38.7568

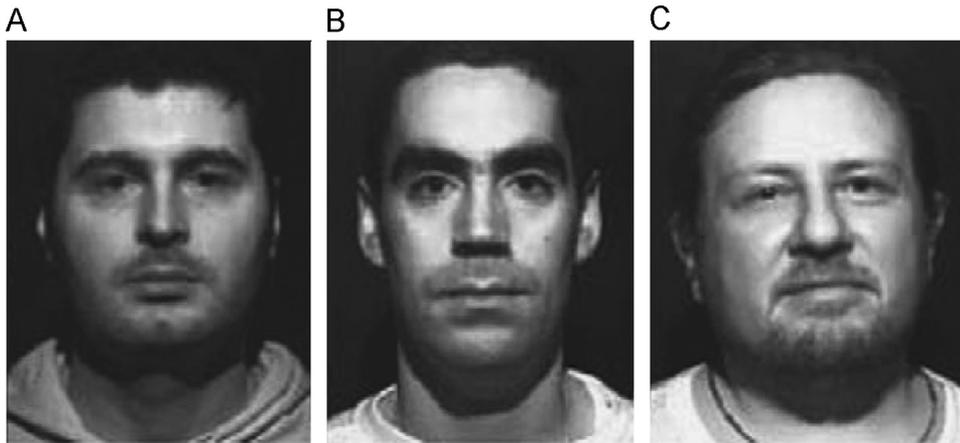


Fig. 10. Real faces used for experiments.

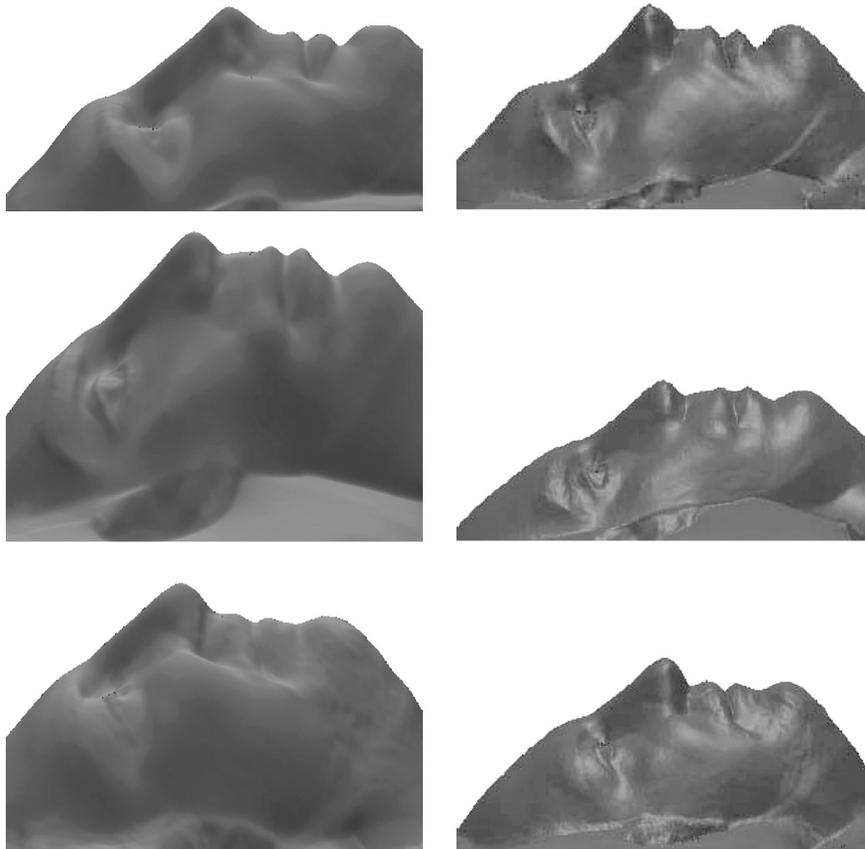


Fig. 11. The profile view of the obtained 3D surfaces using (left) the default illumination directions and (right) the proposed ones. The difference in the obtained details is obvious.

provides more accurate estimates of the optimal illumination direction in terms of optimal normals and albedo estimation keeping the number of lights constant but selecting the proposed illumination directions.

Acknowledgements

This work was supported by the EPSRC project EP/E028659/1 Face Recognition using Photometric Stereo.

We would like to thank also A. Hilton and N. Nadтока from Centre for Vision Speech and Signal Processing of the University of Surrey for supplying the 3D scanner for data collection.

References

- [1] N. Alldrin, S. Mallick, D. Kriegman, Resolving the generalized bas-relief ambiguity by entropy minimization, in: Conference on Computer Vision and Pattern Recognition (CVPR), 2007, pp. 1–7.
- [2] J. Barron, D. Fleet, S. Beauchemin, Performance of optical flow techniques, *International Journal of Computer Vision* 12 (1) (1994) 43–77.
- [3] S. Barsky, M. Petrou, Shadows and highlights detection in 4-source colour photometric stereo, in: International Conference on Image Processing, vol. 3, 2001, pp. 967–970.
- [4] S. Barsky, M. Petrou, The 4-source photometric stereo technique for 3-dimensional surfaces in the presence of highlights and shadows, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (10) (2003) 1239–1252.
- [5] S. Barsky, M. Petrou, Design issues for a colour photometric stereo system, *Journal of Mathematical Imaging and Vision* 24 (1) (2006) 143–162.
- [6] P. Besl, N. McKay, A method for registration of 3-d shapes, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (2) (1992) 239–256.
- [7] E. Candes, J. Romberg, l1-magic: recovery of sparse signals via convex programming, URL: www.acm.caltech.edu/l1magic/downloads/l1magic.pdf, 2007.
- [8] E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory* 52 (2) (2007) 489–509.
- [9] E. Candes, T. Tao, Near optimal signal recovery from random projections: universal encoding strategies? *IEEE Transactions on Information Theory* 52 (2006) 33–61.
- [10] M. Chandraker, S. Agarwal, D. Kriegman, Shadowcuts: photometric stereo with shadows, in: CVPR, 2007.
- [11] M. Chantler, Why illuminant direction is fundamental to texture analysis, *IEE Proceedings, Vision, Image & Signal Processing* 142 (4) (1995) 199–206.
- [12] E. Coleman, R. Jain, Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry, *Computer Vision, Graphics and Image Processing* 18 (1982) 309–328.
- [13] D. Donoho, Compressed sensing, *IEEE Transactions on Information Theory* 52 (4) (2006) 1289–1306.
- [14] O. Drbohlav, M. Chantler, On optimal light configurations in photometric stereo, in: IEEE International Conference on Computer Vision, ICCV, vol. 2, 2005, pp. 1707–1712.
- [15] G. Finlayson, M. Drew, C. Lu, Intrinsic images by entropy minimisation, in: Proceedings of ECCV, 2004, pp. 582–595.
- [16] G. Fyffe, Y. Xueming, P. Debevec, Single-shot, photometric stereo by spectral multiplexing, in: IEEE International Conference on Computational Photography (ICCP), 2011, pp. 1–6.
- [17] A. Georghiadis, Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo, in: 9th International Conference on Computer Vision, vol. 2, 2003, pp. 816–825.
- [18] C. Gullon, Height Recovery of Rough Surfaces from Intensity Images, Ph.D. Thesis, School of Engineering and Physical Sciences, Heriot-Watt University, 2002.
- [19] C. Hernández, G. Vogiatzis, R. Cipolla, Multi-view photometric stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (3) (2008) 548–554.
- [20] A. Hertzmann, S. Seitz, Example-based photometric stereo: shape reconstruction with general, varying brdfs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (8) (2005) 1254–1264.
- [21] B. Horn, Understanding image intensities, *Artificial Intelligence* 8 (11) (1977) 201–231.
- [22] K. Hyeonwoo, B. Wilburn, M. Ben-Ezra, Photometric stereo for dynamic surface orientations, in: ECCV'10 Proceedings of the 11th European Conference on Computer Vision, 2010, pp. 59–72.
- [23] S. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior-point method for large-scale l-regularized least squares, *IEEE Journal of Selected Topics in Signal Processing* 1 (4) (2007) 606–617.
- [24] K. Lee, C. Kuo, Shape reconstruction from photometric stereo, *Computer Vision and Pattern Recognition* (1992) 479–484.
- [25] M. Levine, J. Bhattacharyya, Removing shadows, *Pattern Recognition Letters* 26 (3) (2005) 251–265.
- [26] D. Miyazaki, K. Ikeuchi, Photometric stereo under unknown light sources using robust svd with missing data, in: 17th IEEE International Conference on Image Processing (ICIP), 2010, pp. 4057–4060.
- [27] N. Nadтока, J. Edge, P. Jackson, A. Hilton, Representing dynamics of facial expressions, in: CVMP, 2006.
- [28] S. Nayar, K. Ikeuchi, T. Kanade, Determining shape and reflectance of hybrid surfaces by photometric sampling, *IEEE Transactions on Robotics and Automation* 6 (4) (1990) 418–431.
- [29] H. Ragheb, E. Hancock, A probabilistic framework for specular shape-from-shading, *Pattern Recognition* 36 (2) (2003) 407–427.
- [30] H. Rushmeier, G. Taubin, A. Guezic, Applying shape from lighting variation to bump map capture, in: Eurographics Rendering Workshop, 1997, pp. 35–44.
- [31] F. Sakaue, J. Sato, A new approach of photometric stereo from linear image representation under close lighting, in: IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 2011, pp. 759–766.
- [32] W. Smith, E. Hancock, Estimating cast shadows using sfs and class-based surface completion, in: Proceedings of the 18th International Conference on Pattern Recognition, vol. 4, 2006, pp. 86–90.
- [33] F. Solomon, K. Ikeuchi, Extracting the shape and roughness of specular lobe objects using four light photometric stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (4) (1996) 449–454.
- [34] A. Spence, M. Chantler, Optimal illumination for three-image photometric stereo acquisition of texture, in: Proceedings of the 3rd International Workshop on Texture Analysis and Synthesis, 2003, pp. 89–94.
- [35] A. Spence, M. Chantler, Optimal illumination for three-image photometric stereo using sensitivity analysis, *IEE Proceedings, Vision, Image and Signal Processing* 153 (2) (2006) 149–159.
- [36] J. Sun, M. Smith, L. Smith, S. Midha, J. Bamber, Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities, *Image and Vision Computing* 25 (7) (2007) 1050–1057.
- [37] H. Tagare, J. deFigueiredo, A theory of photometric stereo for a class of diffuse non-lambertian surfaces, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (2) (1991) 133–152.
- [38] V. Argyriou, S. Barsky, M. Petrou, Generalisation of photometric stereo technique to q-illuminants, in: Proceedings of BMVC, 2008.
- [39] V. Argyriou, M. Petrou, Recursive photometric stereo when multiple shadows and highlights are present, in: Proceedings of CVPR, 2008.
- [40] R. Woodham, Photometric stereo: a reflectance map technique for determining surface orientation from image intensities, in: SPIE, vol. 155, 1978, pp. 136–143.
- [41] R. Woodham, Photometric method for determining surface orientation from multiple images, *Optical Engineering* 19 (1) (1980) 139–144.
- [42] J. Wright, A. Yang, A. Ganesh, S. Sastry, Y. Ma, Robust face recognition via sparse representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (2) (2008) 210–227.
- [43] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, Y. Ma, Robust photometric stereo via low-rank matrix completion and recovery, in: Proceedings of the 10th Asian Conference on Computer Vision—Volume Part III, 2011, pp. 703–717.
- [44] Q. Zhang, M. Ye, R. Yang, Y. Matsushita, B. Wilburn, H. Yu, Edge-preserving photometric stereo via depth fusion, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012.